

University of Groningen

Analysing hiker movement patterns using GPS data: Implications for park management

Meijles, E.; de Bakker, M.; Groote, P.; Barske, R.

Published in:
Computers, Environment and Urban Systems

DOI:
[10.1016/j.compenvurbsys.2013.07.005](https://doi.org/10.1016/j.compenvurbsys.2013.07.005)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

Publication date:
2014

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Meijles, E., de Bakker, M., Groote, P., & Barske, R. (2014). Analysing hiker movement patterns using GPS data: Implications for park management. *Computers, Environment and Urban Systems*, 47. <https://doi.org/10.1016/j.compenvurbsys.2013.07.005>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Analysing hiker movement patterns using GPS data: Implications for park management

E.W. Meijles*¹, M. de Bakker¹, P.D. Groote¹, R. Barske²

¹ Faculty of Spatial Sciences
University of Groningen
PO Box 800
9700 AV Groningen, The Netherlands
* Corresponding author.
e-mail: e.w.meijles@rug.nl
tel: + 31 50 3634182
fax: + 31 50 3633901

² State Forestry Commission
North Region
PO Box 333
9700 AH Groningen, The Netherlands

Abstract

In natural areas, park management organisations need to cater for economic, environmental, recreation and social demands and values. However, multifunctional land use also creates conflicts. Increased numbers of people using an area could exceed its ecological carrying capacity. The recreational quality of areas could be negatively impacted by conflicts in recreational behaviour. Therefore, park managers require spatio-temporal data on visitor flows, but there appears to be a shortage of suitable visitor data. If there are data available, these often do not deliver the information required by managers and there is little guidance on appropriate monitoring variables. This paper therefore combines user movement analysis with environmental and ecological factors for natural resource management. Through a case study we describe the entire working process from field data acquisition to usable park management information. GPS and itinerary data from 138 visitors to the Drents-Friese Wold National Park (the Netherlands) were collected to estimate visitor densities and distribution patterns within the park. Data acquisition is efficient in the working process, but careful error handling is a time consuming but necessary part of it. We introduce the definition of ‘hard’ and ‘soft’ errors to make spatial analysis more flexible. We show that walking speed, trip time and spatial distributions varied between park visitor groups. Classification and Regression Tree (CART) analysis showed that factors such as the presence of marked trails, distance to facilities but not land use (such as forest or arable land) explained visitor distributions. Patterns differ between respondent groups based on group size and composition, which is also true for itinerary variables such as walking speed and trip time. The combination of high resolution location data with itinerary information from respondents provides a good impression of the different walking preferences of different respondent groups. We conclude therefore that combining GPS data with itinerary information is a useful tool in profiling different natural park visitors. This is useful information for park managers in steering tourists and in catering for different visitor demands in natural parks.

Keywords

Spatio-temporal analysis, GPS tracking, natural areas, park management

1. Introduction

In natural areas, park management organisations need to cater for economic, environmental, recreational and social demands and values (Geneletti & Van Duren, 2008). Therefore, park management goals not only focus on nature conservation and nature development and redevelopment, but also include recreational targets. However, multifunctional land use also creates conflicts. An increase in the number of people using an area could mean its ecological carrying capacity is exceeded (e.g. Hadwen et al., 2007; Lyon et al., 2011; Wimpey & Marion, 2011). An area's recreational quality could be negatively influenced by conflicts in recreational behaviour (Ligtenberg et al., 2008; Orellana, 2012). Therefore, park managers not only need information on ecological and environmental values, but also require spatio-temporal data on visitor flows. Analysing tourist behaviour – such as places visited, time spent and facility use – can help managers adapt infrastructure and facilities to offer more diverse options to different visitor groups (Dye & Shaw, 2007; Holyoak & Carson, 2009; Wolf et al., 2012) or to route visitors to a range of park locations to avoid overcrowding and to achieve greater matching of visitor and interest (O'Connor et al., 2005; Lyon et al., 2011). Moreover, spatial visitor flow information could be used to define ecological zones and facilitate recreation routing to avoid ecological carrying capacity overload in natural areas (Freuler & Hunziker, 2007; Lyon et al., 2011; Orellana et al., 2012).

There appears to be a lack of sufficiently detailed visitor data for natural areas suitable for supporting park management definitions (Wolf et al., 2012). Where data are available, these often do not deliver the information needed by managers (Hadwen et al., 2007) and there is relatively little guidance on the monitoring variables required (Wolf et al., 2012). This paper therefore combines user movement analysis with environmental and ecological factors for natural resource management. Using a case study, we describe the entire working process from field data acquisition to usable park management information. The novelty of our paper lies primarily in its interdisciplinary character. We bridge a gap between GI technology and methodologies on the one hand, and a real-world problem from a park management perspective on the other. Secondly, special attention is paid to data quality and error handling. We distinguish between 'hard' and 'soft' errors. Thirdly, this paper illustrates the use of the non-parametric Classification and Regression Trees analysis (CART) to assess the visitor density as a function of landscape characteristics. Fourthly, the paper demonstrates the value of itinerary information combined with GPS data in distinguishing between different visitor groups.

In the paper, we first provide a background of GPS tracking in Section 2. In Section 3, we introduce the Drents-Friese Wold National Park as study area. In Section 4, the methodology is outlined describing the full working process from data acquisition via data quality to analysis. It also includes a description of the different datasets used. Section 5 describes results and discussion of the analyses, including hard and soft error handling, differences in walking behaviour between groups and spatial distribution patterns. We finalise the paper by concluding remarks and recommendations on the working process, data quality, spatial movement patterns and the possible use for park managers.

2. GPS tracking

In recent years, (agent-based) computer simulations have been applied to assess the expected visitor distributions in natural areas (e.g. Lawson, 2006). For model validation purposes, empirical field visitor data are required (O'Connor et al., 2005; Van Marwijk, 2009; Orrelana et al., 2012). GPS techniques are widely available to obtain field data. GPS data acquisition is usually fast and yields large, detailed datasets, but the use of these data is often limited to visual (e.g. Taczanowska et al., 2008) or exploratory analysis (e.g. Chen et al., 2011). Several authors have shown that GPS or local spatial transmission/receiving techniques are suitable for further spatial analysis. Their focus is on the method or technique (e.g. Laube & Purves, 2006; Taczanowska et al., 2008), the entire sample (e.g. Ligtenberg et al., 2008), the individual (e.g. Chen et al., 2011), or on separating group behaviour solely based on spatial patterns (O'Connor et al., 2005; Dias et al., 2008). Although the possibility of separating groups based on motives and group composition is mentioned, its application is limited. Some researchers explicitly choose between time-space activity diaries (Xiao-Ting & Bi-Hu, 2012) or GPS data logging (O'Connor et al., 2005), but the literature available on analysis that combines GPS data logging with (qualitative) itinerary information is limited (e.g. Van Marwijk, 2009; Wolf et al., 2012). Itinerary information is used to divide respondents into different demographic or motive groups, but not to statistically test the validity of the group separation itself. In addition, the quality analysis and error handling of GPS data are often limited (Van Marwijk, 2009).

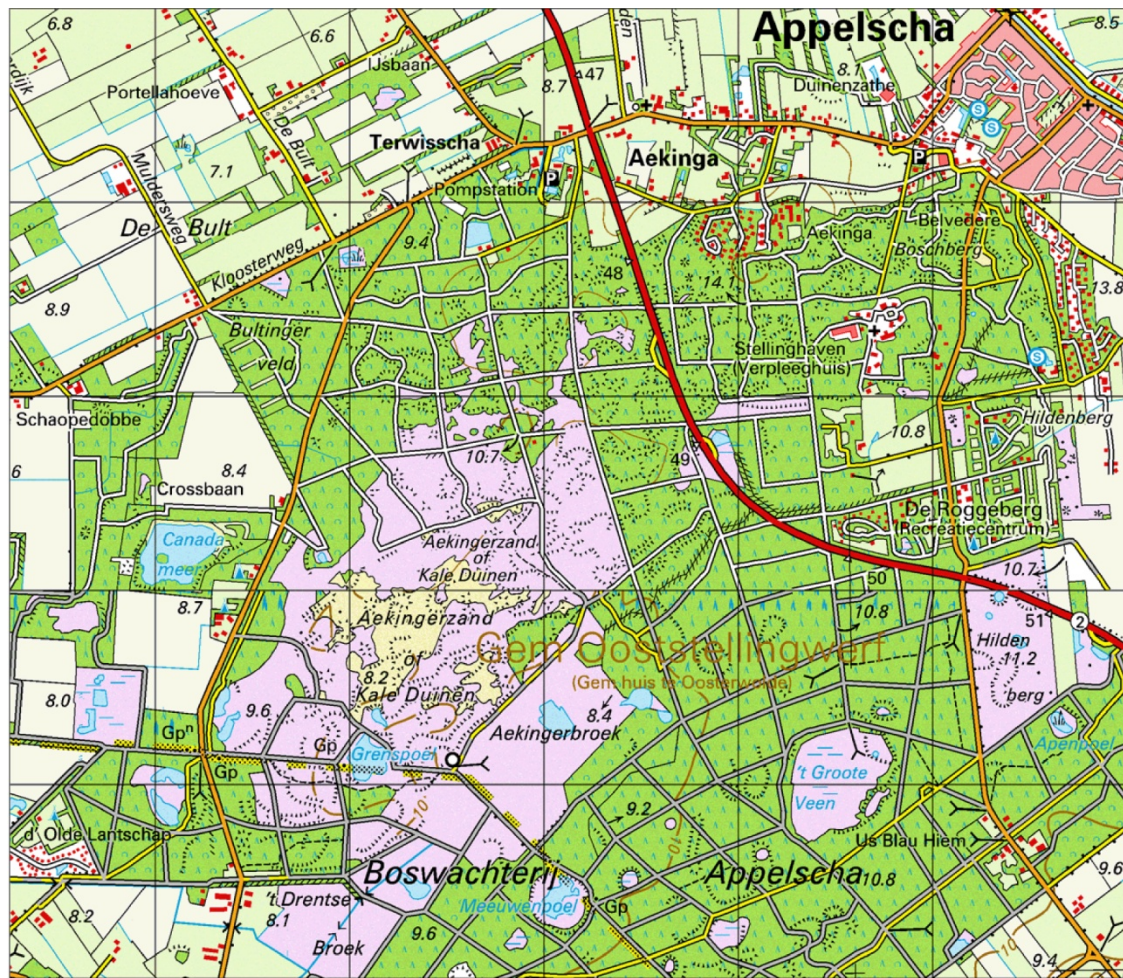
Van Marwijk (2009), Wolf et al. (2012) and Xiao-Ting & Bi-Hu (2012) provide a general overview of the advantages and disadvantages of different techniques for assessing visitor flows in natural areas. These techniques include mental mapping, GPS-logging, space-time diaries, interviews afterwards, anecdotal evidence and direct observation (e.g. following visitors). Since the costs of GPS loggers have decreased significantly in recent years, (O'Connor et al., 2005) quantitative empirical data can now complement the techniques mentioned above. By supplying GPS devices to individual tourists, high resolution data can be obtained (Taczanowska et al., 2008) with little effort from respondents, making data collection less time consuming than other methods (Shoval and Isaacson, 2007). GPS loggers also provide additional information such as duration of stops, speed and off-trail behaviour (Dias et al., 2008; Taczanowska et al., 2008; Wolf et al., 2012).

A disadvantage of the use of GPS loggers is the limited accuracy of signals in densely built-up areas and closed-canopy environments, because a direct line of sight between satellites and the GPS logger is required (Shoval & Isaacson, 2007). In addition, due to the high volume of data – some of it with errors and missing data – data cleaning and analysis is time consuming (Van Marwijk, 2009). A practical drawback is the need for the research participants to start and finish at designated locations so that the GPS loggers can be distributed and collected at the end of the data collection period (Taczanowska et al., 2008). While there are privacy issues associated with this type of research, Taczanowska et al. (2008) have shown that only a few visitors actually refused to take part in a project for this reason. Another point of concern is the extent to which the visitors' awareness of the GPS receiver could influence their walking behaviour (O'Connor et al., 2005). For example, visitors who do not follow rules and regulations are unlikely to participate in a research project, biasing the data obtained (Taczanowska et al., 2008).

3. Drents-Friese Wold National Park

3.1 Case study

In 2011 the managing organisations of the Drents-Friese Wold National Park started updating their park management plans (Elzinga, 2011), which aim to connect the location-specific nature and landscape values with socioeconomic targets for the park. This involves enhancement and restoration of the area's natural and cultural historical values. One aspect of the plans is to create silence and tourist zones. Four different zones with different ecological values and visitor density acceptance have been defined. An estimate of current visitor densities and hiker behaviour is therefore useful as baseline assessment, which makes the National Park a useful area for this study.



 Dutch Land Registry Office, 2006

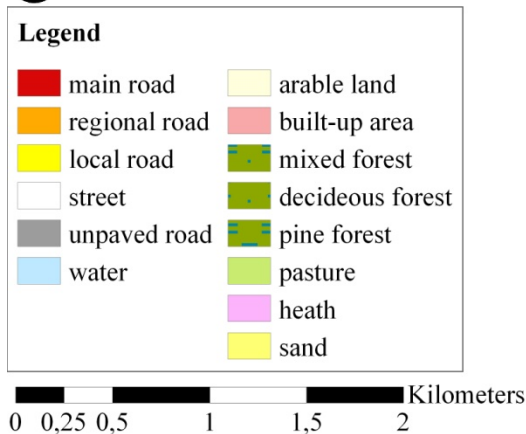


Fig. 1. Map of the study area.

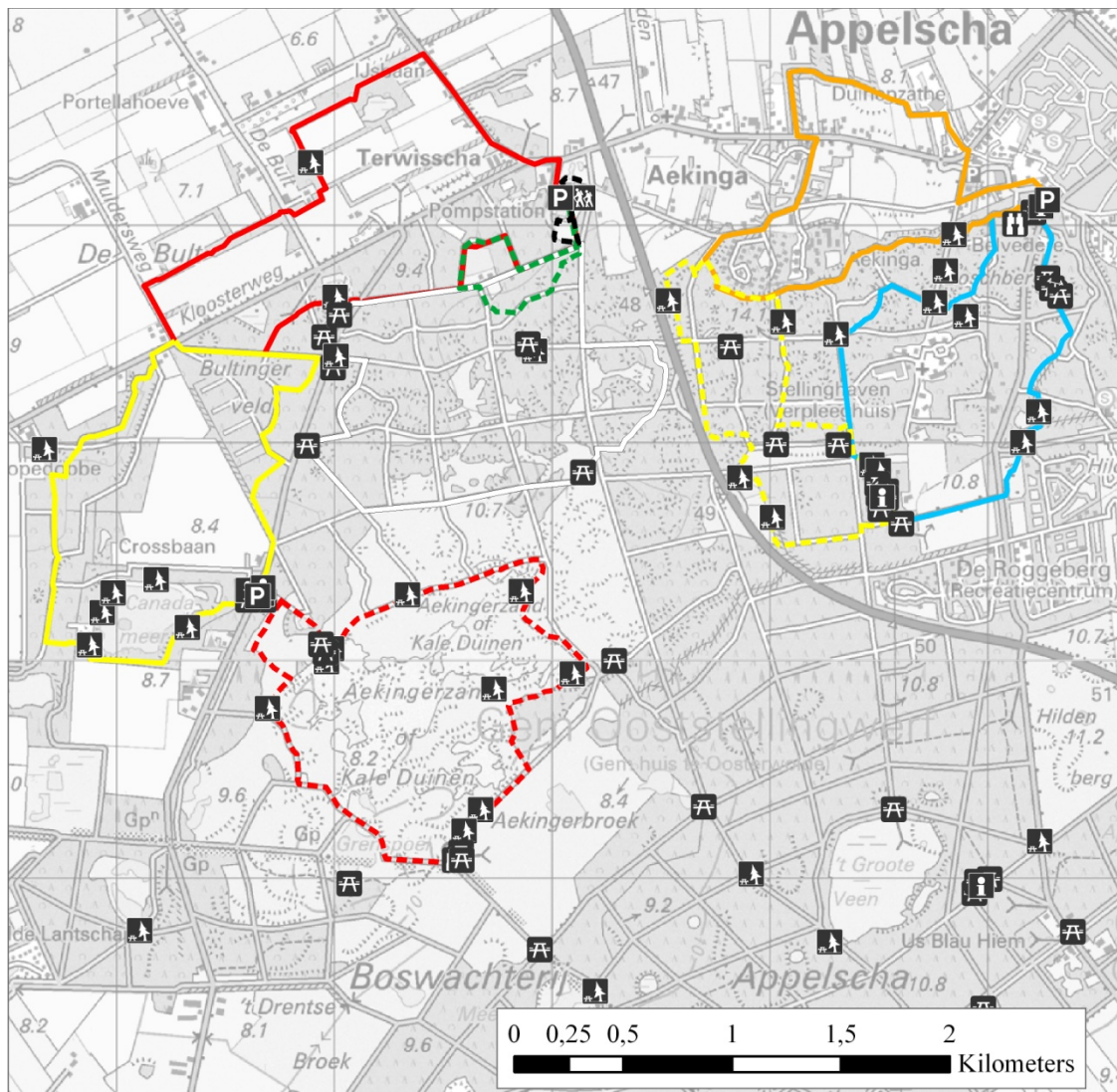
3.2 Study area description


Drents-Friese Wold National Park (6000 ha, Fig. 1) is one of twenty National Parks in the Netherlands and was founded in 2000. The main aims of the park are fourfold (Oranjewoud, 1998): 1. intensification of nature protection and nature redevelopment; 2. promotion of nature-based recreation, 3. stimulation of education and 4. stimulation of research on biotic and abiotic values and recreation in the area.

The managing organisation of the Drents-Friese Wold National Park is a co-operative composed of four bodies: the State Forestry Commission (Dutch: *Staatsbosbeheer*), the National Society of Natural Monuments (*Vereniging Natuurmonumenten*), and the 'Het Drentse Landschap' and 'Maatschappij van Weldadigheid' foundations. The area consists of a landscape mosaic of pine and broad-leaved forests, heaths and open wind-blown sand areas. Biodiversity values are high in the area. The park is also an important recreational area: it is widely used for hiking, cycling and horse riding. Tourist facilities include a visitors' centre focussing on natural values, benches and picnic tables and several signposted trails. The trails range from those accessible to people with physical impairments to rough paths. Most facilities are situated close to car parks, the visitors' centre and signposted trails (see Fig. 2).

The area is zoned based on ecological values (Oranjewoud, 1998). The first zone consists of open, sparsely forested areas, such as heaths, wind-blown sands and brook valleys. These are considered most vulnerable and only limited recreational use is allowed there. Access restrictions are applied during the breeding season, and visitors are informed about the vulnerability of areas to maintain the quietness in the areas. Recreational facilities are sparse.

The second zone is defined as forest edges, bordering the first zone. Limited recreation is allowed, but recreation developments to attract visitors are actively discouraged by the State Forestry Commission. In the third zone, the extensively forested areas, recreational use is allowed. Finally, zone 4 areas are regarded as intensive recreational areas, in which facility quality is optimised for the current visitor numbers.



 Dutch Land Registry Office, 2006

Legend

















	car park		yellow trail (1) - Bultingerveld	- 5.2 km
	visitor centre		yellow trail (2) - Kraaiheidepollen	- 4.7 km
	bench		orange trail - Westeres	- 4.7 km
	picnic table		red trail (1) - De Bult	- 6.0 km
	information panel		red trail (2) - Kale Duinen	- 5.4 km
	play facilities		blue trail - Bosberg	- 4.8 km
	viewpoint		green trail - Family trail	- 2.2 km
			red hats - Gnome trail	- 0.8 km
			white trail - Terwisscha	- 5.3 km

Fig. 2. Visitor facilities and signposted trail network.

4. Methodology

We mapped the work process for this case study (see Figure 3). This provides an impression of how we combined and analysed empirical GPS and survey field data to obtain information usable by park managers. The steps in the figure are described in the following.

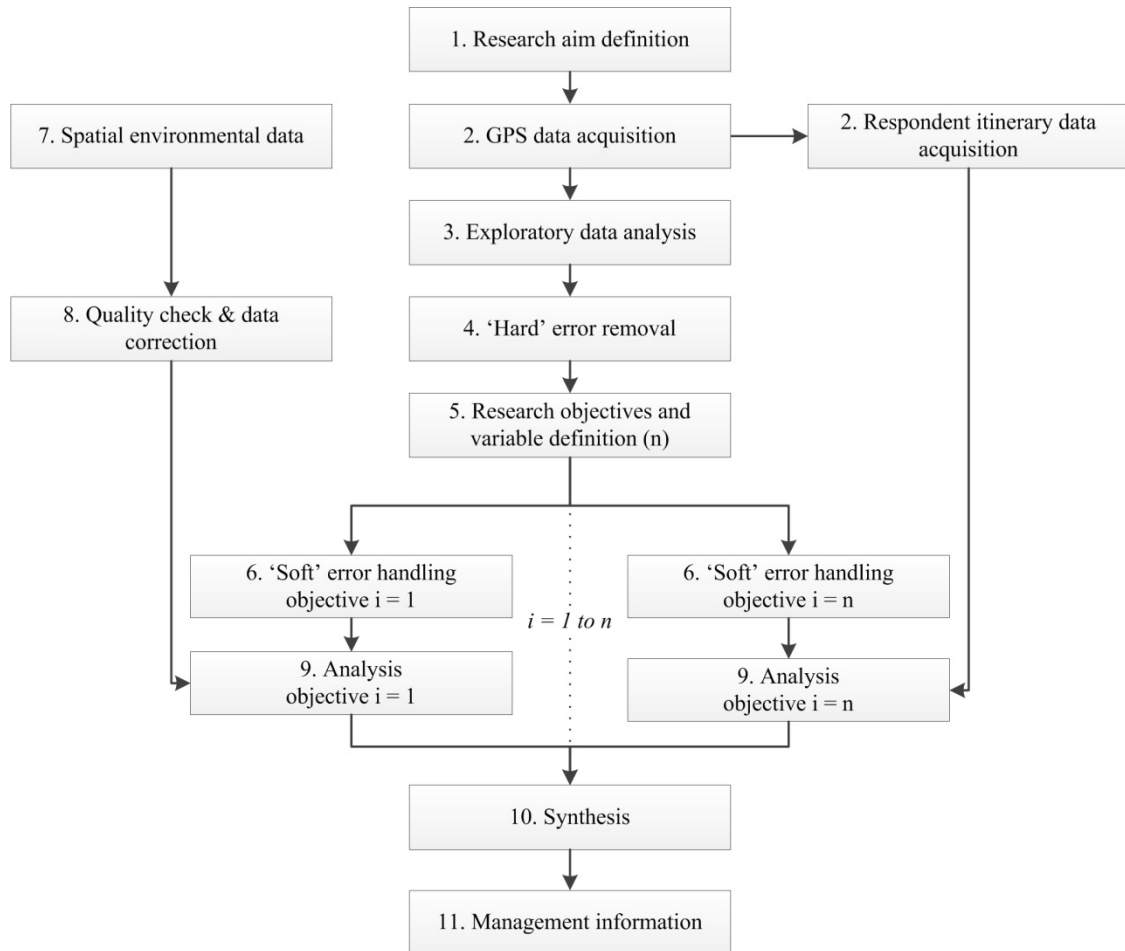


Fig. 3. Schematised working process including the various error handling steps

4.1 Data acquisition

GPS data were collected by students from the University of Groningen (see step 2 in Fig. 3). They randomly asked National Park visitors whether they were willing to participate in the research project. Voluntary respondents were equipped with an Evermore DL-600 GPS logger. During four public holiday and weekend days in May 2010, GPS loggers were supplied at the three car parks (National Park entry points, Fig. 2). The GPS track log was assumed to be representative of the entire group. GPS tracks were recorded on a ten second time interval. On their return to the car park, visitors were asked to complete an anonymous questionnaire to obtain itinerary information, such as group composition, motivations for their visit and the route chosen. These were linked to the data logs based on an identifier number.

4.2 Exploratory data analysis and error handling

The respondents' track logs were entered into a single spatial database. The data were stored in two different formats (points and lines) to cater for different types of analysis. Data were stored as single points and combined with respondent identification. Track logs or geospatial lifelines (observation series of single data points consisting of identification number, location and time; Hornsby & Egenhofer, 2002) were stored as line data per respondent. In an exploratory data analysis (step 3 in the working process) performed to identify patterns, trends and errors (Kundzewicz & Robson, 2004; Dias et al., 2008), individual track logs were visualised using GIS. Identified errors were divided into 'hard' and 'soft' errors. Hard errors were defined as data points that cannot be used in any case and should therefore not be used in any analysis. In our case study, hard errors stored incorrect locations and required removal from the database. We defined soft errors as inconsistencies in the database that do not necessarily cause erroneous results. Soft error types produce inaccurate or incorrect results in specific analyses, but not in others. Hence, the removal of such errors depends on the aim of analysis. Error handling procedures were therefore carried out at several steps (see steps 4 & 6 in Fig. 3) within the working process. Errors were removed by automated algorithms if possible and the remaining errors were deleted by hand (see Sections 5.2 & 5.4).

The theoretical GPS logger accuracy is 15 metres (EverMore Technology, 2007). However, data visualisations of topographic data and aerial photographs in GIS have shown that topographic data also incorporated spatial errors. Furthermore, field tests with the GPS loggers showed that satellite reception was sometimes reduced, especially in forested areas. Therefore, we used an overall accuracy of 25 metres for logged positions, which is higher than the 10 metres employed by Dias et al. (2008) for example.

4.3 Spatial environment data sets

Several spatial datasets were used to relate spatial visitor patterns to environmental factors (see step 7 in Fig. 3). From a visual analysis of aerial photos (Google Earth and BING maps) we concluded that standard topographic data (vector format, scale 1:10,000) from the Dutch Land Registry Office were not sufficiently detailed for analysis in our study, because part of the small-scale infrastructure (walking paths) was not mapped. By combining infrastructure datasets from both the Dutch Land Registry Office and the State Forestry Commission database, a more detailed infrastructure network for the study areas was created. In addition, the wind-blown sand areas had a network of small-scale tracks that were not mapped in the infrastructure datasets. These tracks were therefore digitised using aerial photos. All infrastructure data were converted to lines based on the centre lines of the paths. Parallel paths wholly or partly overlapping in the database were merged to form single paths. The basis of the infrastructure database was the path segment, which was defined as 'a section of the path network between two decision points: that is, a path segment ends at an intersection' (Van Marwijk, 2009, p. 121). The study area infrastructure comprised of 1137 path segments totalling 169 km in length. The path segment formed the basis for further study.

All signposted trails and other tourist facilities such as picnic tables, benches and car parks were retrieved from the State Forestry Commission geodatabase, scale

1:10,000. Land use data based on satellite imagery was retrieved and reclassified from the LGN5 raster database (*Landelijk Grondgebruiksbestand Nederland*; Dutch national land use database, Hazeu, 2005) with a raster cell size of 25 metres.

4.4 Data analysis methodology

After hard data removal, a first visual overview of the spatio-temporal behaviour was obtained. A point density raster map was calculated for an indication of visitor densities. The number of logged points were counted per cell (ten metre resolution) within a radius of 25 metres to compensate for reception inconsistencies (Section 4.2) and converted to the number of points per ha. The resulting map was considered to represent visitor density patterns and to enable visual connections to ecological variables (Lyon et al., 2011) and facilities.

The assessment of differences in walking behaviour across different respondent groups based on demography and respondent group motive was enabled by combining itinerary questionnaire information with the GPS data. On the basis of the questionnaire, respondents were classified into one of four motive groups, and into group size/composition classes, following the work of Van Marwijk (2009). Based on the open questions: ‘Why did you follow this specific route?’ and ‘What is important to you in walking this route?’, the respondents were classified into four motive groups based on combinations of keywords. Although in some cases respondents used the same keywords, combinations with other terms resulted in classification into different groups.

The ‘social and relaxation’ group consisted of people who came to the park for social reasons, where relaxing with friends and family was most important. The natural surroundings were usually a backdrop of their walking trip. People responded with keywords such as ‘family’, ‘social’, ‘children’, ‘fun’, ‘drinking coffee’, ‘being together’ and ‘relaxing’. People in the ‘social and activities’ group also described their reasons from a social point of view (keywords: ‘social’, ‘family’ or ‘children’), but specifically included keywords such as ‘playing in the forest’, ‘exploring on the children’s trail’, ‘entertainment’ and ‘experience’ in their responses. People in the ‘nature and rest’ group were looking for rest and the pleasure of natural areas, and mentioned terms such as ‘nature’, ‘peace and quiet’, ‘freedom’ or ‘off the beaten track’. The ‘walking as exercise’ group included people who liked walking because of the physical activity itself. They used typical combinations of keywords such as ‘long distance’, ‘moving’, ‘walking the dog’ and ‘jogging’.

In addition, the respondents were also classified on the basis of the size of the groups they were in (one or two people, three to five people, and six people or more) and whether they accompanied young children (0-6 years).

We defined walking behaviour using the following variables: speed (walking pace), total walking distance, trip time and off-path behaviour. Visitor densities on a path were expected to be dependent on the signposted trails, environmental factors (landscape composition) and the available facilities.

In Dutch culture, nature is generally regarded as something fragile that should be protected and not entered. Wandering off paths in natural areas is therefore not generally accepted. It was therefore expected that people tended not to wander off the paths. In this case study we assumed that locations over 25 metres from infrastructure could be considered as being off the paths, based on theoretical GPS logger error margins, topographic data precision and field reception (Section 4.2).

Following Van Marwijk (2009), it was assumed that visitor density is a function of environmental variables, network layout and visitor facilities. A person's presence at a specific location at time t is not independent of his or her location at time $t-1$. In other words, walkers are constrained in their route choice by a limited set of path segments at a specified junction and therefore visitor densities in a path network are not spatially independent. This has already been noted by Hägerstrand (1970). To relate environmental variables to visitor densities, we adopted the non-parametric classification and regression tree (CART) methodology, which creates a prediction model represented as a tree (Loh, 2008). If the dependent variable is categorical, this method produces a classification tree, while a continuous dependent variable will result in a regression tree (Loh, 2008; Wilkinson, 1992). The trees are formed by a set of rules based either on categorical or numerical independent variables. A Chi-squared Automatic Interaction Detection algorithm (CHAID: Kass, 1980) looks for the most statistically significant explanatory variable. This variable is used to split a node into two or more branches and then the process is repeated for each child node. All branches of the tree end in a terminal leaf, in which all observations fall into a single leaf (Wilkinson, 1992).

Within this study's framework, we regarded the dependent variable 'point density' as a measure of respondent visits, with the path segment as the sampling unit. The more points on the segment, the more often the segment was visited, or the longer individual respondents stayed on the path. We defined point density as the number of logged points per 100 metres of path segment.

The number of logged points was counted by a buffer operation for each path segment. All the waypoints within a 25 metre radius from a path segment (corresponding to the error margin, see Section 4) were assigned to that specific segment. As 283 segments were shorter than the error margin of 25 metres (often due to spatial database inconsistencies, such as non-perfect path crossings) these were left out of the analysis. In addition, path segments closer than 50 metres to the car parks were removed, as their point density scores were heavily influenced by people preparing or coming back from their walks, and were therefore regarded as not being representative for of walking trip.

The regression tree analysis was carried out for the points logged for the entire population (138 respondents, 1137 path segments). As an example of studying the effect of different group compositions on point density values in the network, we included a CART analysis of groups with young children up to 6 years old (a sub-sample of 51 of 138 respondents, 1137 path segments). We also used the proportion of respondents with children as the dependent variable in the CART analysis, which we regarded as a measure of the 'attractiveness' of path segments. The ratio between the points logged by respondents with children and all logged points was calculated for each path segment. Only path segments with logged points were included ($n=807$), as division by zero is impossible.

Independent continuous variables were distance to facilities such as car parks, viewpoints, information panels, playing facilities and picnic tables or park benches. Distances were calculated via the path network. Land use type was a nominal variable. Path segments that were part of a signposted route were regarded as binary explanatory variables. All variables were entered into the CART analyses. The branching limit was defined by the minimum size of the child node (25 path segment) or after a maximum of three levels of branching.

5. Results and discussion

5.1 Data acquisition

A total of 138 respondents participated in the research. In keeping with the findings of Taczanowska et al. (2008), only a few people worried about possible privacy issues. Our experience was therefore, that the limiting factor in obtaining a large sample was the number of GPS loggers available rather than the number of participants willing to volunteer in the research. Some respondents walked alone, but most were in groups. In total, nearly 55,000 single data points were recorded. Each data point consisted of a set of geographical coordinates, the time and a respondent identification number. All respondents completed the questionnaire after their trip, although three questionnaires (2%) left some missing answers.

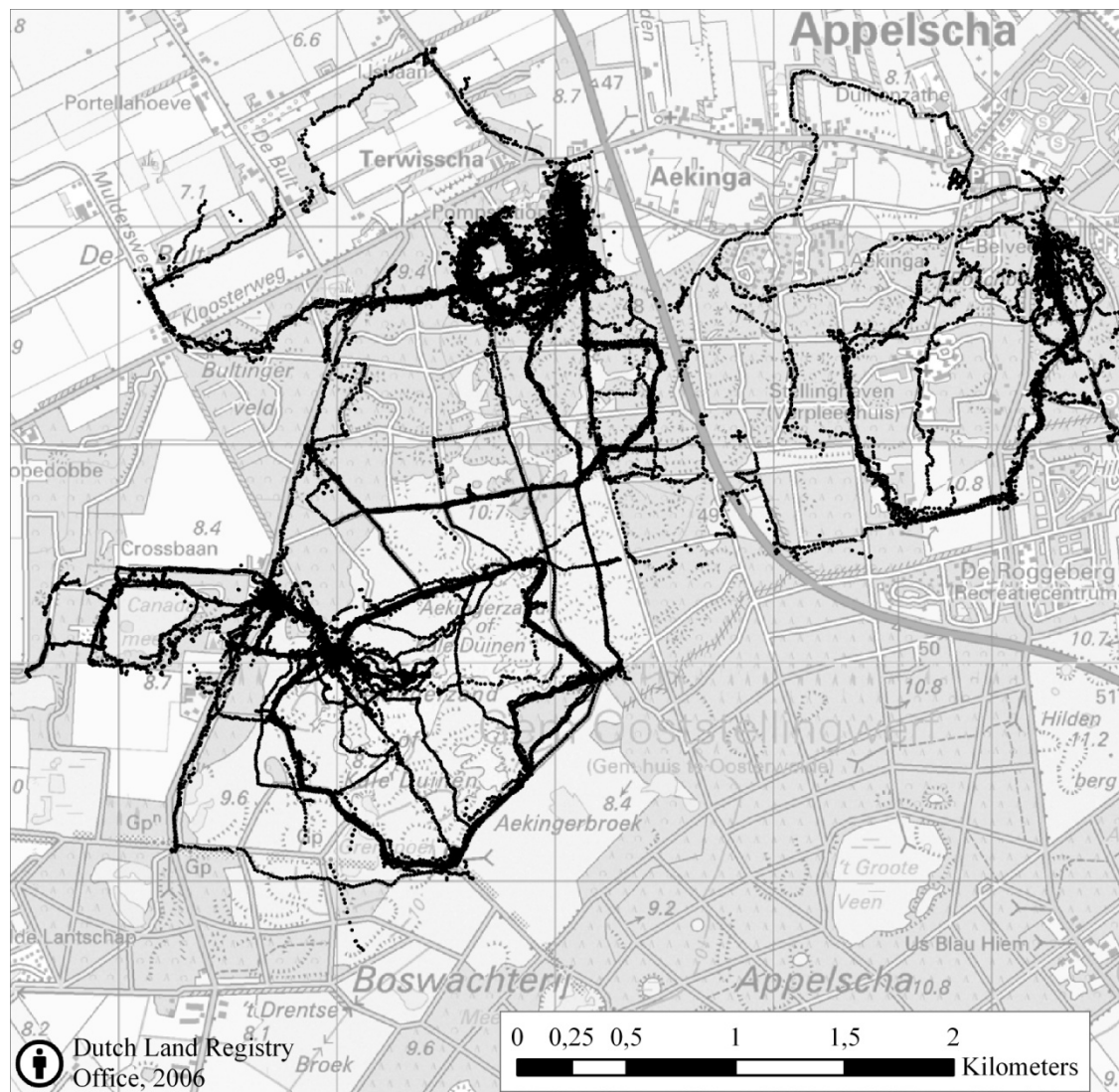


Fig. 4. Uncorrected GPS point data on the topographical map.

5.2 Exploratory data analysis and hard error removal

Fig. 4 shows the spatial distribution of the uncorrected GPS data points acquired before error handling. Some preliminary spatial patterns can be interpreted, such as large numbers of data points near car parks and the identification of the most and least frequented paths. Possible hard errors (see Section 4.2) can also be identified, such as suddenly ending track logs and outliers. Such errors can often be attributed to closed canopy conditions in natural areas resulting in satellite signal loss (Taczanowska et al., 2008). In our case, bad satellite reception could also be due to loggers being kept in bags or pockets by the respondents. Error identification and handling was treated at different points within the working process (see steps 4 & 6 in Fig. 3). The hard error removal is discussed here; soft error handling is described in section 5.4.

In the first step of the hard error analysis the completeness of the track logs was checked. From a total of 138 respondents (totalling 54,941 individual logged points), there were five instances where the GPS logger only recorded a small fraction of the route followed (see Table 1 for details). These entire track logs (error type 1, see Table 1) were removed from the database.

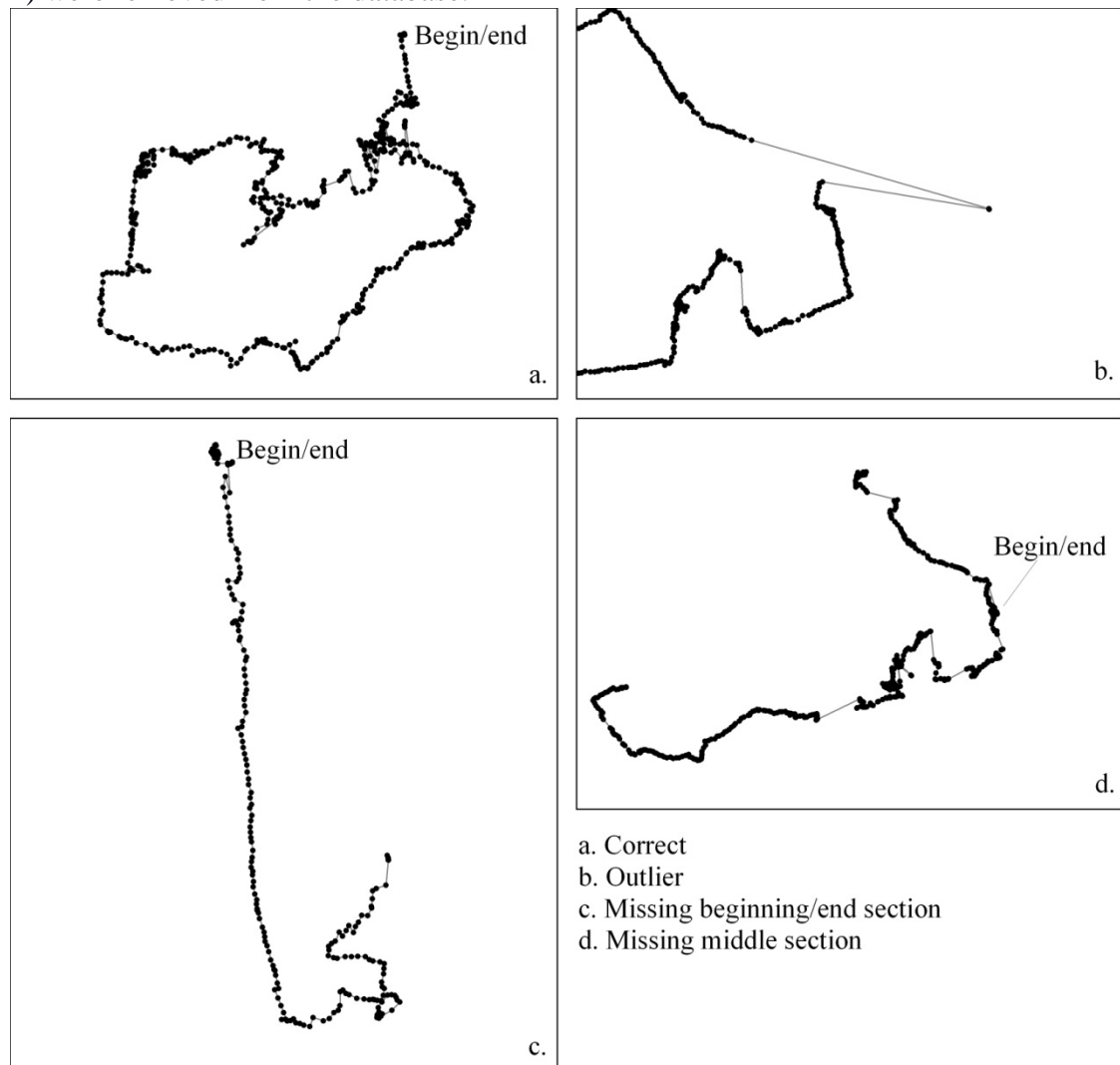


Fig. 5. Examples of spatial errors.

Secondly, a visual inspection for possible location errors was carried out. A number of outliers (referred to as ‘drift’ by Xiao-Ting & Bi-Hu, 2012) were identified: logged points that were distinctively separated from their predecessors and/or successors in

the track log (see Fig. 5b). Using an iterative algorithm based on the assumptions that walking speeds between points are unlikely to be higher than running speed (25 km/h), outliers were deleted from the database (error type 2, see Table 1). We allowed a higher maximum speed than the one used by Dias et al. (2008) (7 km/h). That cut-off speed seemed unrealistically low to us, as running children were included in the sample, for instance. A visual analysis later showed some remaining outliers, which were removed by hand. At this stage, 3.8% of the logged points (see Table 1) were regarded as hard errors and were removed from the database.

Table 1. Error treatment of point data.

Error type	Visual analysis	Error treatment	N	%
Total number of logged points:			54,941	100
1: Fragment ('hard')	Very short section	Logs with - lengths less than 200 hundred metres and/or - only covering the car park and/or - a low number of logged points, generally fewer than 100 were removed from the database.	1,530	2.8
2: Location error ('hard')	Outlier or drift	If the distance between consecutive points was exceptionally long, the speed (distance divided by logging frequency: 10 s) should be relatively high. Speeds higher than running speed (i.e. 25 km/h) are unlikely. As outliers were often not individual but occurred in clusters, points and their associated speeds were iteratively removed.	563	1.0
3: Temporal error ('soft')	Begin/endpoint not the same	All track logs (lines) were visually inspected and errors were noted in a quality field in the geodatabase. Error acceptability (see main text) depends on the analysis and therefore track logs were selected on the basis of the quality field.	see Table 2	
	Middle section missing: straight line with points far apart	All track logs were visually inspected and errors were noted in a quality field in the geodatabase. Error acceptability (see main text) depends on the analysis and therefore track logs were selected on the basis of the quality field.	see Table 2	

5.3 Point density visualisation

Fig. 6 is an illustration of a point density analysis. Visual analysis shows a clear correspondence with the signposted trail infrastructure and with visitor facilities (see Fig. 2), which corresponds to observations by Wolf et al. (2012) that there are higher visitor concentrations on a small number of paths. Density values tend to be high near car parks, confirming that the distance to parking facilities is an important factor for hiker routes (cf. Van Marwijk, 2009). Other high point densities are the visitor centre, the signposted trails (especially the designated family and children's trails) and the wind-blown sand area. On the other hand, there are many paths that do not seem to be

used. These are generally the non-signposted paths. Most visitors appear to walk through natural forest, heaths and sandy areas, whereas arable land and pasture are rarely visited (see Figs. 1 & 6).

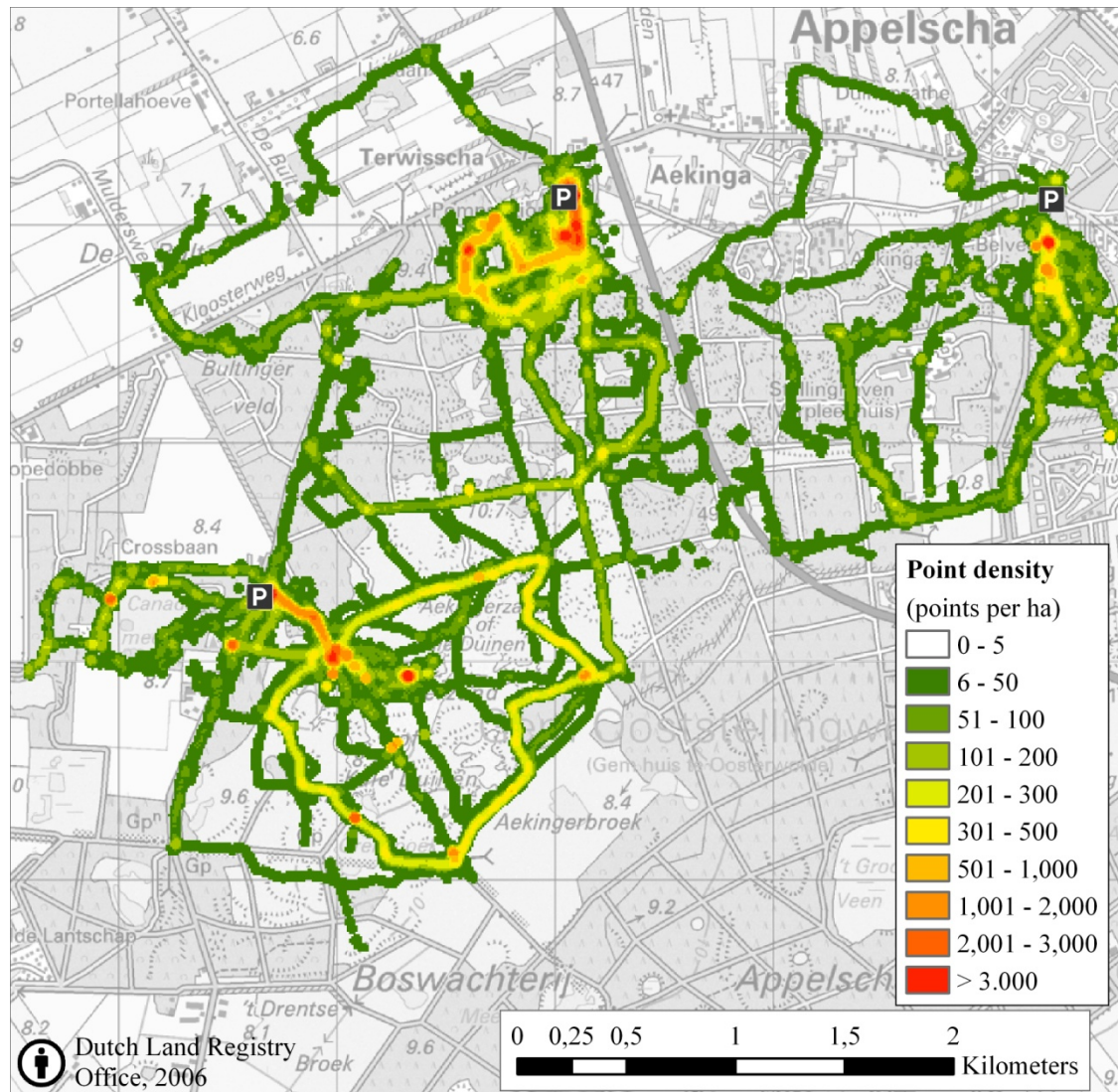


Fig.6. Point density map of visitor GPS data logs using a 25 metre search radius.

5.4 Soft error handling

Some GPS loggers did not record the entire track. In these cases, track logs had missing sections, such as the beginning or end section (see Fig. 5c) or parts in the middle (see Fig. 5d). We identified these as ‘soft’ errors (see Section 4.2), as the possible use of these data points depended on the type of analysis being performed. For example, track logs with missing data cannot be used for trip distance calculations. But as long as the data missing from that track log are not substantial, data can be used for average speed analyses. Of the 133 track logs, only 53 (around 40%) were entirely complete and therefore usable for total length calculations. Taczanowska et al. (2008) and Van Marwijk (2009) have also noted the relatively large amounts of missing data in their GPS track logs. In their case studies they found that 59% and 78% of the track

logs, respectively, were of sufficient quality to derive all important individual routing variables, and the rest were not.

When establishing an idea of average walking speed per respondent, however, a certain amount of missing data was acceptable. It was assumed that if more than 60% of each individual track log was complete, the log could be used for average speed calculations. Track completeness was calculated by analysing the number of logged points (logging frequency of ten seconds) over the total trip time. Of all logs, 111 track logs (83%) were suitable for further analysis.

There are road matching algorithms available for spatial distribution analysis, such as simple shortest distance calculations, Kalman filters, fuzzy rules and others (Quddus et al., 2003; Jagadeesh et al., 2004; Tradišauskas et al., 2009) to complete missing sections of the route followed. Obviously, these generated locations cannot be used for off-path analyses and are of limited use for speed calculations. The path network was relatively dense in our case study, leaving several interpretations for different routes. To avoid the introduction to of possible errors, road matching was not employed.

For trip time calculations, a large error was acceptable. As long as the beginning and end times were recorded, the trip time can be calculated. Therefore, logs with missing middle sections could also be used, yielding a suitability of 56% (77 logs; see Table 2).

Table 2. Track log error overview.

Error description	n	%
1. Fragment	5	3.6%
2. Short section	19	13.8%
3. Begin or end section missing	32	23.1%
4. Middle section missing	24	17.4%
5. Combination of 3 and 4	5	3.6%
6. Complete track log	53	38.4%
Total number of track logs	138	100%

5.5 Temporal and distance variables

Walking speeds differed significantly ($p < 0.05$; see Table 3) between respondent groups. Visitors from the ‘walking as exercise’ group moved more quickly on average than those who visited the national park to socialise. The slowest group consisted of those who engaged in activities, which may be explained by defining ‘activities’ as playing outdoor games. The trip time and total walking distance did not differ significantly between groups, although there was a tendency for the ‘nature and rest’ group to walk longer distances. The difference in sample sizes can be attributed to soft error definitions (Section 5.4).

Table 3. Differences in route variables between motive groups.

	Speed (km/h)			Distance (km)			Trip time (mins)		
	n	mean	s.d.	n	mean	s.d.	n	mean	s.d.
Social and relaxation	46	3.3	0.88	22	4.7	2.48	36	91	36
Social and activities	20	2.8	0.92	12	4.1	1.58	16	90	44
Nature and rest	38	3.4	0.86	17	5.5	3.20	21	100	58
Walking as exercise	7	4.1	0.66	1	4.1		3	108	26

Total	111	3.3	0.91	52	5.0	2.62	76	94	44
F		<i>4.626</i>			0.710			0.301	
Sig.		<i>0.004</i>			0.551			0.824	

Based on ANOVA. Statistically significant results ($p < 0.05$) in italics.

The observed behaviour corresponds with the work by O'Connor et al. (2005), showing that trip times are relatively uniform between respondents. This can be explained by the generally flat paths in the area, which provide good access for all ages and abilities. We expected people to remain in the area for more or less the same period, but to cover longer or shorter distances as average speeds would differ. The data does not support this hypothesis, however.

Larger groups and groups with children walk shorter distances and at a slower pace ($p < 0.005$; see Table 4). Hike durations were significantly different between different group compositions (roughly between one to two hours), and can be attributed to the tendency for groups with children to stay for less time (67 minutes) than other groups (98-107 minutes). It can therefore be concluded that different group compositions without children generally stayed in the area for the same period, but because of different average walking speeds between groups, longer or shorter distances were covered. This is in contrast to motive groups, where the differences were only significant with respect to average walking speed.

Table 4. Differences in route variables between group compositions.

Group composition	Speed (km/h)			Distance (km)			Trip time (mins)		
	n	mean	s.d.	n	mean	s.d.	n	mean	s.d.
Alone/with partner	37	3.7	0.92	13	6.7	3.17	19	107	53
3-5 people, all aged over 6	27	3.3	0.88	11	5.5	2.10	20	98	39
3-5 people including children aged 0-6	26	3.1	0.84	16	3.3	1.81	20	67	28
More than 5 people, with/without children	21	2.8	0.64	13	5.0	1.93	18	103	43
Total	111	3.3	0.90	53	4.8	2.58	77	94	44
F		<i>5.680</i>			<i>5.578</i>			<i>3.904</i>	
Sig.		<i>0.001</i>			<i>0.002</i>			<i>0.012</i>	

Based on ANOVA. Statistically significant results ($p < 0.05$) in italics.

5.6 Off-path behaviour

Table 5 shows that the off-path behaviour of the respondents was limited, as the percentage of points traced off the paths was generally less than 2%. Few respondents roamed freely off the tracks in the heaths, pastures and arable fields. The wind-blown sand areas, however, were an exception, with nearly 45% of the log points off the paths. Excluding the sand dunes, our figures are lower than for example Dias et al. (2008), as 31% of their respondents wandered off the paths. Freuler and Hunziker's (2007) research in a snow-covered landscape found a high percentage (68%) wandering off the paths into surrounding protected areas. However, the paths may have been less visible to visitors due to the snowy conditions. The low percentage of off-path walking presented in our study could also be due to the types of visitors

studied. Other types, such as outdoor hikers and mountain-bikers could display different off-path behaviour, but these were not included in our research project. It is remarkable that the table also shows data points in open water (shown in italics). As it is unlikely that people walked through open water, this suggests data quality issues with the GPS data, or the land use and/or path segment databases.

Table 5. Number of occasions respondents were recorded off the paths.

Land use type	Total (n)	On (n)	Off (n)	%
Built-up areas	1,255	1,246	9	0.7%
Heaths and other open areas	13,921	13,720	201	1.4%
Pasture	7,365	7,240	125	1.7%
Arable fields	362	355	7	1.9%
<i>Water</i>	<i>1,531</i>	<i>1,489</i>	<i>42</i>	<i>2.7%</i>
Forest	21,585	20,680	905	4.2%
Wind-blown sands	6,791	3,765	3,026	44.6%
Total	52,848	48,533	4,315	100%

5.7 Relating visitor's spatial behaviour to environmental characteristics

We related the visitors' spatio-temporal behaviour to the characteristics of the landscape and tourist facilities, in which visit numbers are expressed as visitor density values per 100-metre path segments (point density, section 4.4) using CART. Fig. 7 shows the regression tree of the CART analysis, Fig. 8 shows the resulting spatial distribution of the different CART classes.

The distance of the path segment from the car park ($p < 0.001$) is the first regression tree branching point and is therefore the most significant explanatory variable. The point density shows a decreasing trend with distance classes up to 1,693 metres from the car park, which means that respondents tended to stay relatively near to the car park. The second branching point shows that respondents prefer marked trails for most distance classes, as path segments on signposted routes show relatively high point density values. However, for path segments between 348 and 1,000 metres from the car park, the distance to a park bench or picnic table is a more important branching factor, though this relationship is not linear. Intermediate distances to a seating facility show relatively low point densities, but the point density is significantly higher both nearer and further away ($p < 0.001$).

At the third branching level, the explaining variables become more diverse, including distance to play facilities ($p < 0.005$), park benches or picnic tables ($p < 0.001$) or in heath areas ($p < 0.001$). There is a high degree in variability of point densities at all branching levels, as shown by the high standard deviations, but branching points are still significant ($p < 0.01$).

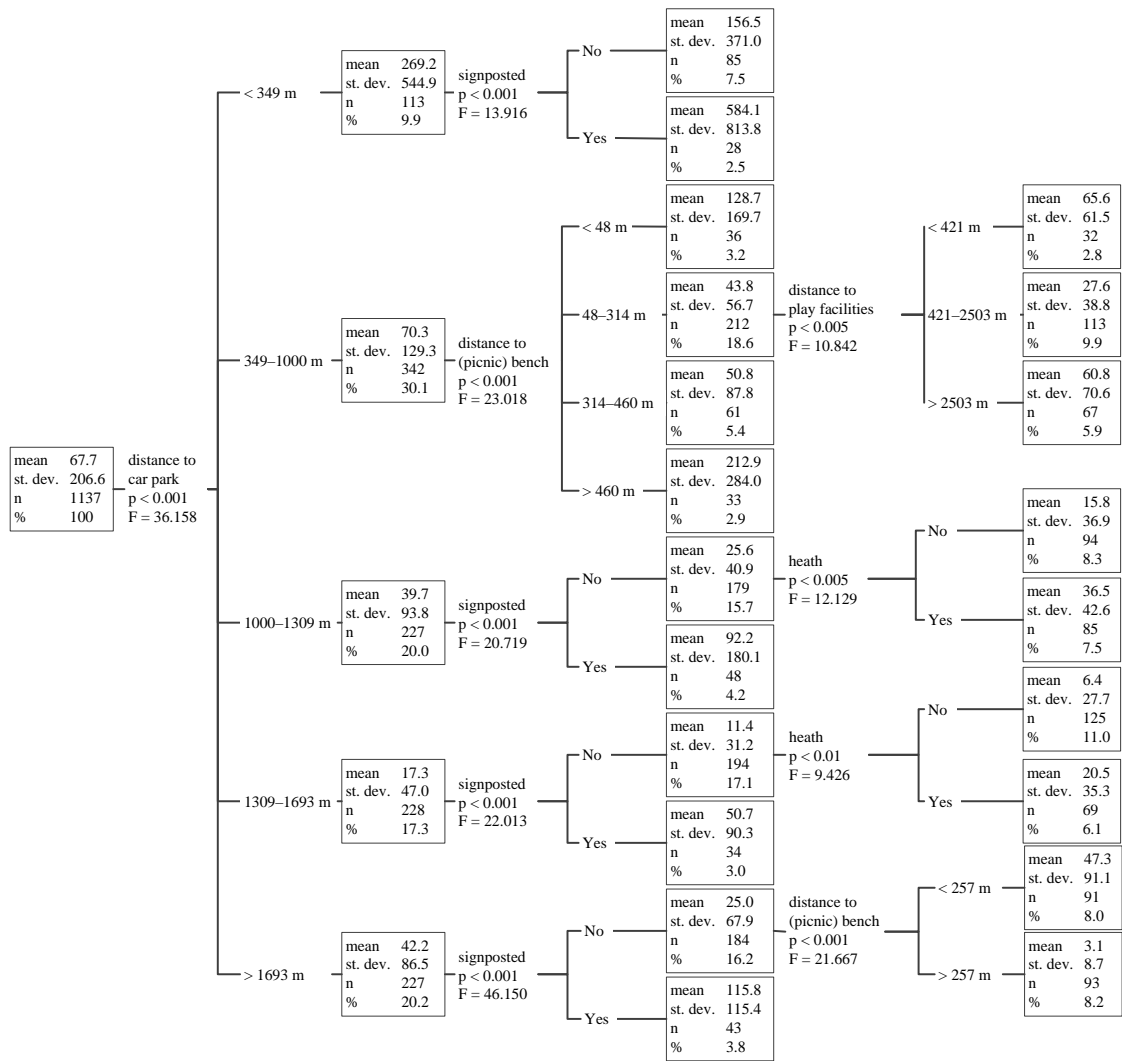
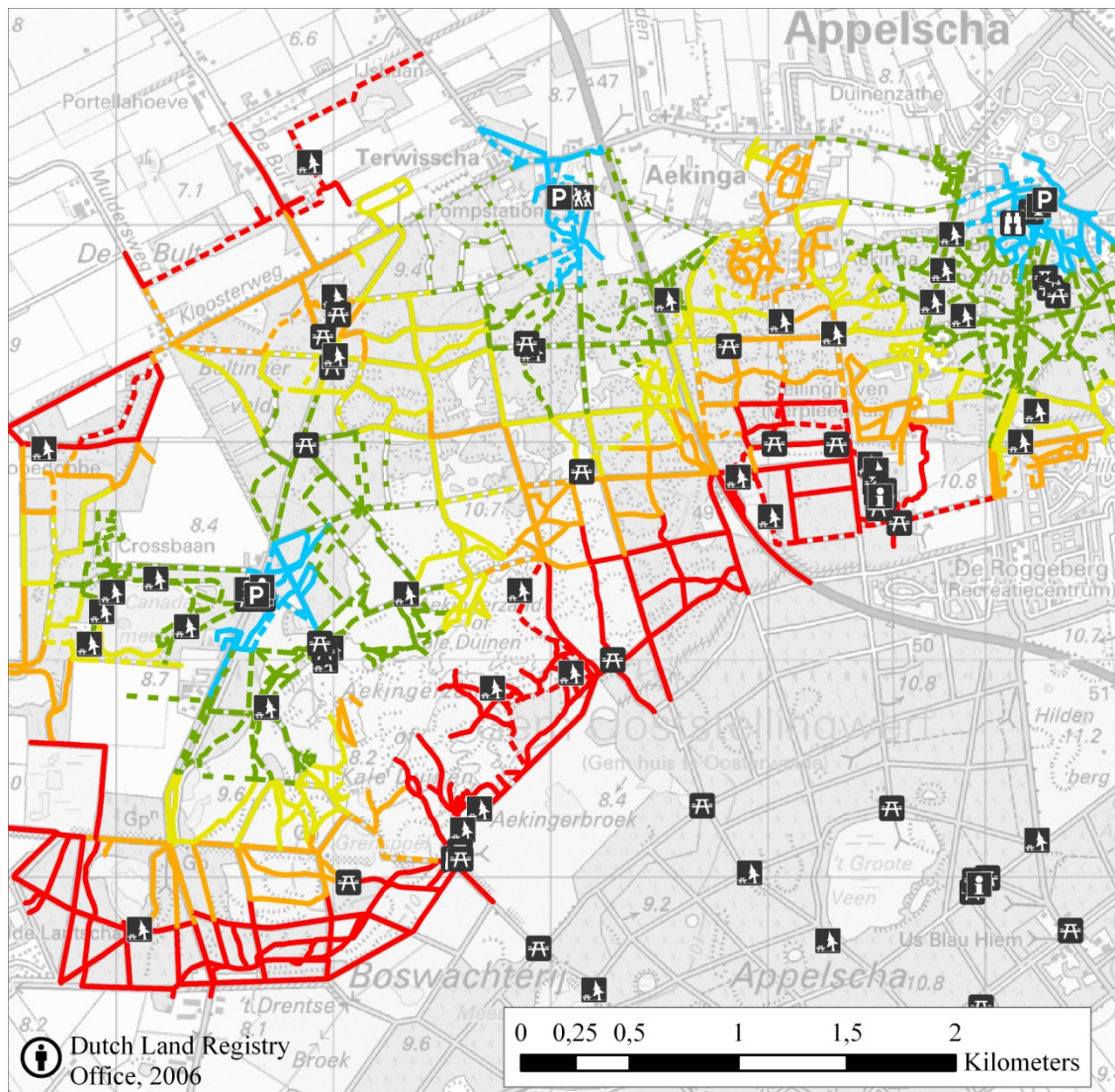


Fig.7. CART results for number of points per 100-metre path segment.



Legend	CART classification	Distance to car park (m)	Signposted Route	Distance to (picnic) bench (m)
P car park	—	< 349	No	
visitor centre	- - -	< 349	Yes	
bench	—	349-1000		< 48
picnic table	- - -	349-1000		48-314
information panel	- . - .	349-1000		314-460
play facilities	349-1000		> 460
viewpoint	—	1000-1309	No	
	- - -	1000-1309	Yes	
	—	1309-1693	No	
	- - -	1309-1693	Yes	
	—	> 1693	No	
	- - -	> 1693	Yes	

Fig.8. Geographical distribution of CART results for number of points per 100-metre path segment. For cartographical reasons, only the second branching level is shown.

The sub-sample of respondents with young children (Fig. 9) also indicates that the distance to the car park is the most significant explaining variable, as indicated by the first branch point ($p < 0.001$). Although the distance classes slightly deviate from the entire respondent population, again there is a general decreasing trend of point density with distance. At the second branch point, at distances shorter than 1000 metres from the car park, results are very similar to the full sample. However, the pattern is much more diverse than the entire sample. Up to 1,969 metres, the distance to play facilities becomes the most significant explanatory variable ($p < 0.001$), within which it is remarkable that point densities tend to be higher further away from the play facilities. This is in contrast to expectations, as these facilities are created specifically for this respondent group. Finally, respondents with young children who travel further from the car park ($> 1,969$ metres) prefer heaths.

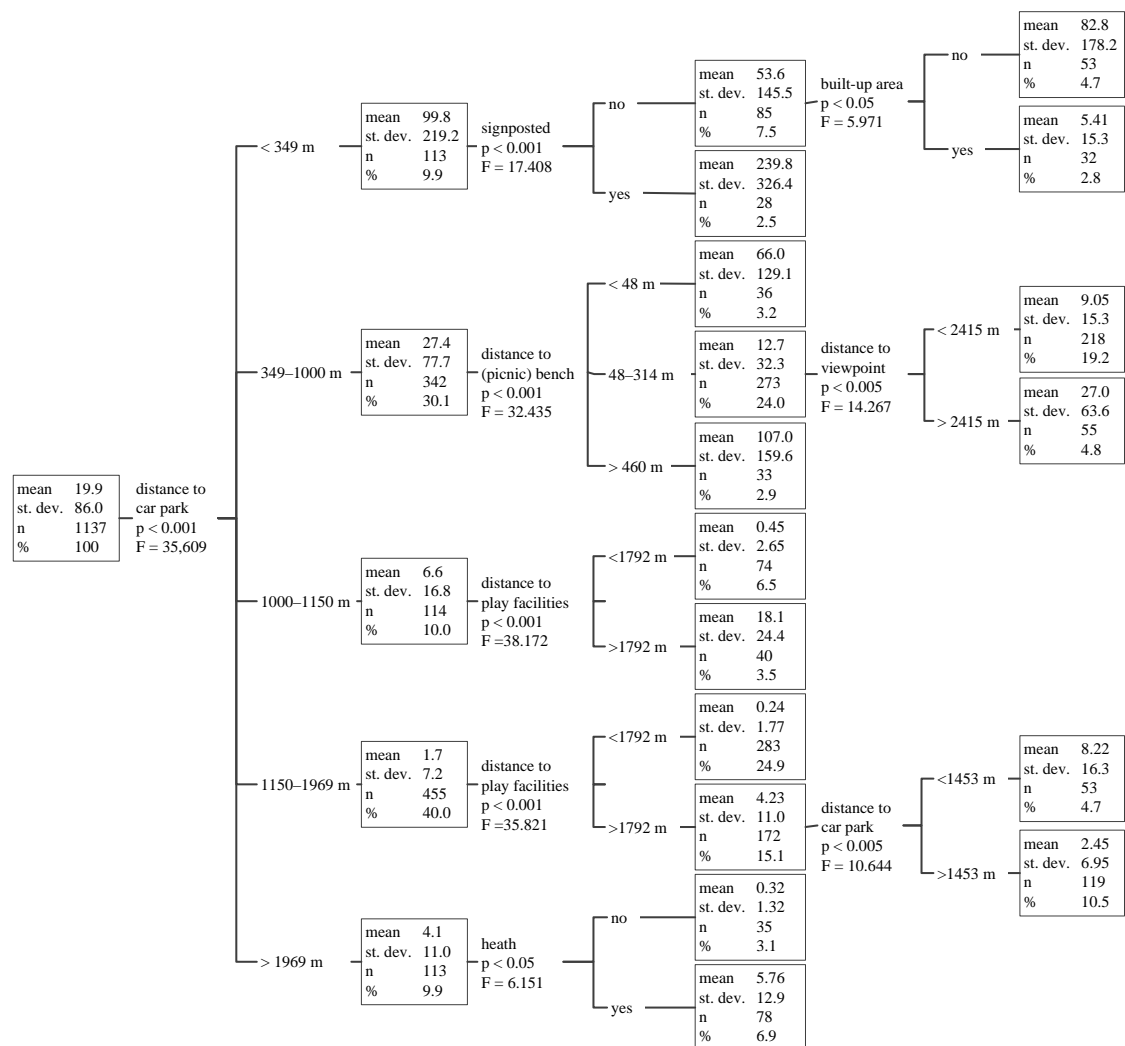


Fig.9. Regression tree of point densities of respondents with children aged 0-6.

When using the proportion of people with young children per path segment as the dependent variable (Fig. 10), the distance to play facilities is the most significant explanatory variable. On paths closer than 348 metres from a playground, an average of 34.9% of the respondents were with young children. This decreases to around 1.2% at 812-1,899 metres. It is striking that at longer distances, the proportion increases again. In fact, paths close to a playground and close to a viewpoint tend to attract a

relatively higher proportion of people with children, which is also true for paths far from playing facilities and further away from viewpoints. This could mean that the grouping variable ‘with young children’ actually incorporates several subgroups. A possible reason for this is that people with babies and toddlers in prams are not confined to short distances, but are still classified into the same group as children that walk themselves. These factors could also be included in future research.

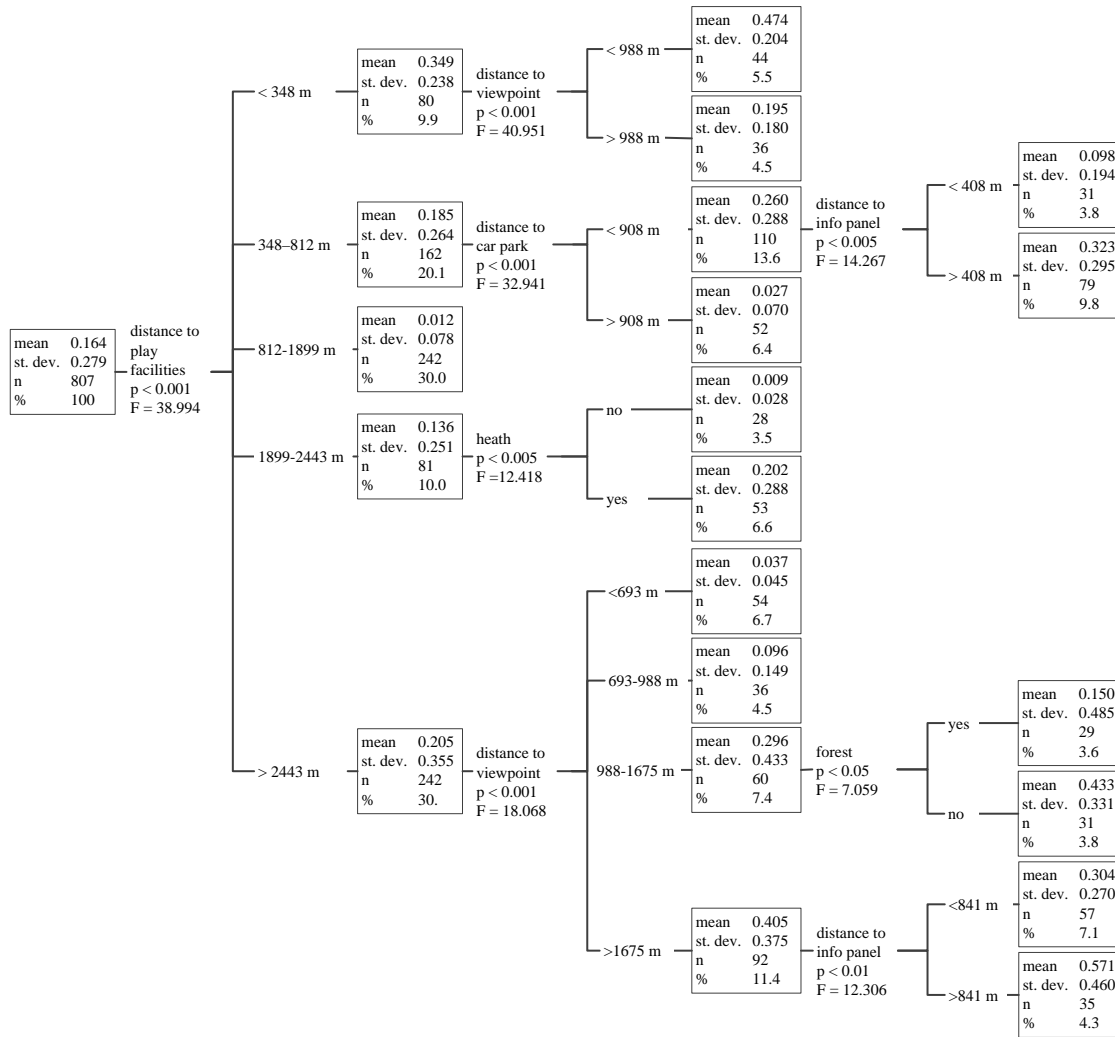


Fig.10. Regression tree for the proportion of respondents with children aged 0-6.

From the analyses we conclude that paths close to car parks are used more frequently than others and this explanatory variable is most significant. At a lower level, but still highly significant, signposted trails show higher visitor numbers than paths not part of a trail. It is remarkable, however, that there is no tendency for specific trails to be followed. In other words, there is no favourite individual signposted trail (Fig. 2).

Other distance variables such as benches, picnic tables and play facilities significantly affect point densities, but are ambiguous, as their impact varies between different distance from the car park classes and specific subgroups of the respondent population. Moreover, no significant relationship between point densities and distance to information panels has been established, which can be interpreted as meaning that

panels do not attract many people, or that people tend not to slow down (thus logging more points) to read them. However, it could be that people know the area and therefore do not need the information. We have not combined visit frequency in this analysis. The distance to a viewpoint plays little role in the regression trees for the entire population, but does play a role for respondents with young children. It is striking that the various land use types – apart from heath and forest – make no significant contributions to point densities, which contradicts earlier visual analyses that found that the wind-blown sand area can be regarded as hotspots. Apparently, the variables mentioned above are much more important than land use for the respondents in this study.

6. Conclusions and recommendations

6.1 Working process and data quality

We conclude from this case study that few problems are involved in the collection of GPS and itinerary data. Visitors are generally willing to carry a GPS logger during their stay in a park and they do not regard this as a violation of their privacy. It is a relatively quick method to obtain spatio-temporal visitor data. The use of smartphones with GPS devices and tracking apps (storing journeys on the mobile phones or directly transmitting them to an online data logging database) can further improve data acquisition, although privacy concerns may become greater if visitors are asked to use their own device.

Although it proved to be easy to collect large amounts of individual space-time data, these were not without errors. Data consolidation and error handling turned out to be time intensive. The definition of general rules and effective algorithms for error handling needed to be carefully thought through. In our experience, visual analysis proved to be crucial, although time-consuming, in the data consolidation process. The distinction between ‘hard’ and ‘soft’ errors proved useful, as each variable in the analysis could be sensitive to different error types.

The results have shown that spatial database quality is an important issue. Present-day GPS loggers are generally very accurate (typically within 10-15 metres), despite possible reception constraints in forested and built-up areas, and are reliable in terms of logging data over time. This also means that in nature-based research, where high resolution GPS data is combined with environmental data derived from other spatial databases, the latter also need to be of high resolution (i.e. 10 metre or 1:1000 scale). In our case, small paths not present in the infrastructure databases may have led to the incorrect conclusion that people were wandering off the paths.

In addition, standard, readily available datasets need to be used with caution. For instance, we found that it is doubtful that hikers change their walking patterns to favour particular land use areas, but probably value their surroundings in terms of the openness or variability of the landscape mosaic (Coeterier, 1994; Tveit et al., 2006). Landscape appreciation datasets, such as openness (cf. Weitkamp et al., 2011) could therefore be more useful than land use databases.

6.2 Hiker movements

It can be concluded from the point density visualisation and the regression tree analysis that the majority of respondents stayed relatively close to car parks and follow

the signposted trails in the National Park. This could indicate that routing of visitors is effective, and that once visitors are on a trail, they stay on the trail. Freuling and Hunziker (2007) and Dias et al. (2008) have shown that visitor behaviour can be changed by providing visitors with information on the trail, points of interest and the sensitivity of wildlife to disturbance. In such cases, the number of visitors tending to stay on the signposted trails increased. This is important for park managers, as they can focus on the access to trails, rather than on keeping visitors on them, which can be an important tool in ecological zoning. However, our results show that only a limited number of visitors are reached, as information panels do not seem to attract visitors. Extending the research to include the frequency of visits or familiarity with the area is needed to further draw conclusions from this.

We do not yet know whether people tend to simply follow signposts, or follow their customary weekly walk, which coincides with the signposted trails. A follow-up study that we are currently involved in includes deliberately altering some of the signposted trails to study how this influences peoples' route decisions. A possible follow-up could be to include sequential spatio-temporal analysis to study patterns in activities or location visits over time. Such analyses could help park managers optimise facility use.

When defining different respondent groups on the basis of motive or group composition, we found that groups differed in hiking behaviour in terms of walking speeds, distances and environmental variables but the staying time (except for groups with young children) did not differ between groups. This was despite the assumed homogeneity of the sample (weekend tourists visiting the national park by car and therefore starting and finishing at the same location). We showed in an example that including itinerary information in the regression trees can provide more detailed information on the walking patterns of subgroups classified by group composition. In this example, we showed that the proportion of people with young children statistically differs between path segments and can be related to environmental variables. The regression method itself finds statistical differences between groups, rather than those defined by the researcher. Extending this method by using motive groups in regression tree analyses could also yield more insight into spatial distribution patterns. Park management organisations could use this information to create 'visitor profiles' to design new signposted trails and further improve different trails for different demands. Visitor profiles could possibly also be linked to different ecological zones in the area.

As our dataset only included weekend tourists, it is not a representative sample of natural park visitors. To obtain a better understanding of visitor patterns in an area, cyclists, mountain bikers, horse riders and joggers could also be included. Furthermore, people not entering and leaving the park from the same location (such as long distance walkers) were not incorporated in our sample. The extent to which the track logs are representative of the group is also debatable: the walking pattern of individuals within the group could differ. Off-track walking could have been underestimated, for instance for groups where the adult is carrying the GPS logger and playing children regularly wander off the paths. In our study, the number of people walking off-path was very low, apart from in sandy areas. However, park managers are often worried about off-path behaviour because of its presumed disturbing effect on the ecology. Taczanowska et al. (2008) suggest that the GPS receivers are most likely carried by people who do not intend to violate regulations, so the results could be biased towards, for example, on-path behaviour. A specific recommendation to park managers could be to study the spatial pattern and off-path behaviour of visitors who are more tempted to go off-path, such as geocachers. Hidden caches in the area could

reveal hotspots or, in the case of multi-caches, suggest informal routes that our point density visualisation does not show.

Representativity also plays a role in the itinerary data, as the reason for the hike could differ from person to person. Combining research with individual respondents, such as that done by Orellana et al. (2012) and Dias et al. (2008), with group analysis methods such as those introduced by Laube and Purves (2006), could improve representativeness, and will also provide insight into group dynamics.

With this paper, we have described a case study in which easily obtainable GPS data and itinerary trip information was translated to information usable by park managers of a specific natural area. This paper is therefore an example that bridges the gap between technological methodologies and the practical problems natural parks face. We have shown that careful error handling is necessary to filter out errors. Environmental data should be used critically when relating spatial distribution patterns to environmental variables determining walking behaviour. Integrating itinerary information with GPS data proved a useful tool in profiling different natural park visitors, as shown in our regression tree analysis. This is useful both for steering tourists in the natural areas in combination with ecological zoning, and for catering to different visitor demands.

Acknowledgements

The Netherlands State Forestry Commission is gratefully acknowledged for their support. We thank the students J. de Boer, M. Daggenvoorde, A. van Eckand and L. Huttenhuis, who assisted in data acquisition and preliminary analysis. Discussions on statistics with Prof. J. Tinbergen and Dr. W. Meester were very fruitful. The assistance of E. Uibel and T. Koelman in solving technical problems is greatly appreciated. We are grateful for feedback from the three anonymous reviewers.

References

- Chen, J., Shaw, S. L., Yu, H., Lu, F., Chai, Y. & Jia, Q. (2011). Exploratory data analysis of activity diary data: a space–time GIS approach. *Journal of Transport Geography* 19, 394-404.
- Coeterier, J. F. (1996). Dominant attributes in the perception and evaluation of the Dutch landscape. *Landscape and Urban Planning* 34, 27-44.
- Dias, E. S., Edwardes, A. J. & Purves, R. S. (2008). Analysing and aggregating visitor tracks in a protected area In: Stein A., Shi, W., & Bijker W. (Eds) *Quality Aspects in Spatial Data Mining* (pp. 265-282). CRC Press.
- Dye, A. S. & Shaw, S. L. (2007). A GIS-based spatial decision support system for tourists of Great Smoky Mountains National Park. *Journal of Retailing and Consumer Services*, 14, 269-278.
- Elzinga, G. (2011) *Plan van aanpak Beheer en Inrichtingsplan Drents-Friese Wold*. <http://www.np-drentsfriesewold.nl/documents/documents/plan-van-aanpak-bip-dfw-versie-13-september-2011.pdf>. Last accessed on 12 December 2012.
- Ever More Technology, 2007. *DL-600 Data Logger User's Manual*. Hsinchu, Taiwan. 9p.
- Freuler, B. & Hunziker, M. (2007). Recreation activities in protected areas: bridging the gap between the attitudes and behaviour of snowshoe walkers. *Forest Snow and Landscape Research*, 81 (1/2), 191-206.

- Geneletti, D. & Van Duren, I. (2008). Protected area zoning for conservation and use: A combination of spatial multicriteria and multiobjective evaluation. *Landscape and Urban planning*, 85, 97-110.
- Hadwen, W. L., Hill, W. & Pickering, C. M. (2007). Icons under threat: Why monitoring visitors and their ecological impacts in protected areas matters. *Ecological Management & Restoration*, 8 (3) 177-181.
- Hägerstrand, T. (1970). What about people in regional science? *Papers in Regional Science* 24, (1) 7-24.
- Hazeu, G.W. (2005). *Landelijk Grondgebruiksbestand Nederland (LGN5): Vervaardiging, nauwkeurigheid en gebruik*. Alterra report 1213, Wageningen, 92p.
- Holyoak, N. & Carson, D. (2009). Modelling self-drive tourist travel patterns in desert Australasian. In: *Proceedings of the 32nd Australasian transport research forum (ATRF) meeting*, Auckland, Australia.
- Hornsby, K. & Egenhofer, M. J. (2002). Modeling moving objects over multiple granularities. *Annals of Mathematics and Artificial Intelligence* 36, 177-194.
- G. R. Jagadeesh, G. R., Srikanthan, T. & Zhang, X. D. (2004). A Map Matching Method for GPS-Based Real-Time Vehicle Location. *The Journal of Navigation*, 57, 429-440.
- Kass, G.V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, 29, 119-127.
- Kundzewicz, Z. W. & Robson, A. J. (2004). Change detection in hydrological records: a review of the methodology. *Hydrological Sciences Journal*, 49, (1) 7-19.
- Laube, P. & Purves, R. S. (2006). An approach to evaluating motion pattern detection techniques in spatio-temporal data. *Computers, Environment and Urban Systems*, 30(3), 347-374.
- Lawson, S.R. (2006). Computer Simulation as a Tool for Planning and Management of Visitor Use in Protected Natural Areas. *Journal of Sustainable Tourism*, 14 (6), 600-617.
- Ligtenberg, A., Van Marwijk, R. B. M., Moelans, B. & Kuijpers, B. (2008). Recognizing patterns of movements in visitor flows in nature areas. In: *Proceedings Fourth International Conference on Monitoring and management of visitor flows in recreational and protected areas*, Montecatini Terme, Italy.
- Lyon, K., Cottrell, S.P., Siikamaki, P. & Van Marwijk, R.B.M. (2011). Biodiversity Hotspots and Visitor Flows in Oulanka National Park, Finland. *Scandinavian Journal of Hospitality and Tourism*, 11, 100-111
- Loh, W.Y. (2008). Classification and regression trees. In: Ruggeri, F., Kennet, R.S. & Faltin, F.W. (eds). *Encyclopedia of Statistics in Quality and Reliability* (pp. 315-323), John Wiley & Sons.
- O'Connor, A., Zenger, A. & Itami, B. (2005). Geo-temporal tracking and analysis of tourist movement. *Mathematics and computers in Simulation*, 69, 135-150.
- Oranjewoud (1998). *Beheer- en inrichtingsplan Drents-Friese Wold*. 124p. Assen, The Netherlands.
- Orellana, D., Bregt, A. K., Ligtenberg, A. & Wachowicz, M. (2012). Exploring visitor movement patterns in natural recreational areas. *Tourism Management*, 33, 672-682.
- Quddus, M. A., Ochieng, W. Y., Zhao, L. & Noland, R. B. (2003). A general map matching algorithm for transport telematics applications. *GPS Solutions*, 7, 157-167.

- Shoval, N. & Isaacson, M. (2007) Tracking tourists in the digital age. *Annals of Tourism Research*, 34(1), 141-159.
- Taczanowska, K., Muhar A. & Brandenburg, C. (2008). Potential and limitations of GPS tracking for monitoring spatial and temporal aspects of visitor behaviour in recreational areas. In: Raschi A. & Trampetti S. (eds). *Management for Protection and Sustainable Development* (pp. 451-455). Consiglio Nazionale della Ricerche, Montecatini, Italy.
- N. Tradišauskas, N., Juhl, J., Lahrmann, H. & Jensen, C.S. (2009). Map matching for intelligent speed adaptation. *Intelligent Transport Systems*, 3 (1), 57-66.
- Tveit, M., Ode, A. & Fry, G. (2006). Key concepts in a framework for analysing visual landscape character. *Landscape Research*, 31, 229-255.
- Van Marwijk, R. B. M. (2009). *These routes are made for walking: understanding the transactions between nature, recreational behaviour and environmental meanings in Dwingelderveld National Park, the Netherlands*. PhD thesis Wageningen University, 260p.
- Weitkamp, G., Bregt, A.K. & Van Lammeren, R. (2011). Measuring Visible Space to Assess Landscape Openness. *Landscape Research*, 36(2), 127-150.
- Wilkinson, L. (1992). Tree Structured Data Analysis: AID, CHAID and CART. Sawtooth/SYSTAT Join Software Conference, Evanston, 10p.
- Wimpey, J. & Marion, J. L. (2011). A spatial exploration of informal trail networks within Great Falls Park, VA. *Journal of Environmental Management* 92, 1012-1022.
- Wolf, I. D., Hagenloh, G. & Croft, D. B. (2012). Visitor monitoring along roads and hiking trails: how to determine usage levels in tourist sites. *Tourism Management* 33, 16-28.
- Xiao-Ting, H. & Bi-Hu, W. (2012). Intra-attraction Tourist Spatial-Temporal Behaviour Patterns. *Tourism Geographies*, 14 (4), 625-645.